Heavy tailed priors: an alternative to non-informative priors in the estimation of proportions on small areas

Jairo A. Fúquene P. *
Brenda Betancourt Canizales †

July 15, 2011

Abstract

The behavior of objective robust Bayesian methods in survey sampling is qualitatively different than the traditional noninformative and conjugate Bayesian methods and arguably much more reasonable and acceptable for practitioners and agencies. We explore in this work to use the Cauchy and a new heavy tailed prior proposed by Fúquene, Perez & Pericchi (2011) for binary data in the exponential family to estimate proportions in small areas. The objective robust Bayesian approach is more effective than the traditional case of noninformative or conjugate priors for the estimation of proportions in small areas because when there is a conflict between prior information and the auxiliary information, within or between the small areas, the objective robust priors become noninformative priors and in this sense the prior information is discounted. In order to illustrate the objective robust Bayesian approach, we apply this methodology in a popular example with two types of outliers. Finally, we recommend to use the Cauchy prior in absence or presence of outliers within the small area, and the Fúquene et al. (2011) prior when the outlier is a small area.

Keywords: *Survey Sampling, Exponential Family, Objective Robust Priors, Small Areas Estimation.*

1 Introduction

Little & Zheng (2007) make a comprehensive Bayesian proposal in survey sampling an important field where the Bayesian methods are hardly used. These authors consider in their paper noninformative priors for Bayesian methods in survey sampling settings. We believe that the objective robust Bayesian approach in survey sampling settings could be more effective than the choice of noninformative priors, as suggested by the authors, in order to eliminate antipathy towards methods that involve subjective elements or assumptions. We can address this by

*Department of Applied Mathematics and Statistics Jack Baskin School of Engineering University of California, Baskin Engineering, Room 144, Santa Cruz , USA. jfuquene@ucsc.edu.
†Department of Applied Mathematics and Statistics Jack Baskin School of Engineering University of California, Baskin Engineering, Room 144, Santa Cruz , USA. bbetanc1@soe.ucsc.edu

using objective Bayesian robust priors that are dominated for the likelihood when prior and likelihood information are in conflict. On the other hand, there are recent advances and new proposal for Objective Robust Bayesian Analysis. We use here some of the most recent literature with these contributions. We can find the first in Fúquene, Cook & Pericchi (2009) where Cauchy and Berger's robust heavy-tailed priors are considered and several mathematical results are presented such as the Generalized Polynomial Theorem for robust priors. Also, we use the Student-t-Beta2(1,1,1,$\boldsymbol{\beta}$) heavy tailed prior proposed by Fúquene et al. (2011) for modelling outliers and structural breaks in dynamic linear models. The Fúquene et al. (2011) prior is a new heavy tailed prior founded as the marginal for the location parameter of: a Cauchy density for the location parameter and a Scaled Beta2 (see Jonhnson, Kotz & Balakrishnan (1995)) prior for the square scale.

The Bayesian approach for estimating parameters defined on small areas have had considerable attention in recent years. For example, in the popular book of Rao (2003) there are different proposals for various small area estimation methods using Bayes and hierarchical Bayes methods. On the other hand, the estimation of proportions is one of the most important topics in small areas estimation methods since binary data is often present in survey sampling. Little & Zheng (1980) propose an empirical Bayes approach to estimate small areas using mixed logistic regression models. Stroud (1991) develops a general hierarchical Bayes methodology for univariate natural exponential families. Stroud (1994) shows a proposal for the treatment of binary data for different designs of survey sampling such as simple random, stratified and cluster and two stage sampling. Jiang & Lahiri (2001) propose a frequentist alternative to the hierarchical Bayes methods for the small areas estimation with binary data. However, there is no proposal on using the exponential family for the Binomial likelihood with objective Bayesian robust priors; therefore, our proposal is quite distinct from the previous proposals and it is a novel proposal to the best of our knowledge. We support our proposal by different reasons:

1. Heavy tailed priors are useful in the posterior inference in small areas not only when there is conflict between auxiliary variable (prior information) and sample within the small areas but also when there is conflict between the small areas.

2. The use of prior information with robust priors could be more acceptable for both practitioners in survey sampling and agencies because robust priors discount the prior information when the auxiliary variable (prior information) in the small areas is in conflict with the actual data. Therefore, agencies could see these methods like objective and without "prior" biases.

3. MCMC simulation with robust priors for the estimation of proportions in small areas is fairly simple and it could definitely help the diffusion of robust Bayesian methods in survey sampling for practitioners. In fact, we use in this paper two friendly R-packages that people involved in survey sampling can use easily.

This paper is organized as follows: in Section 2, we give a background of the objective robust Bayesian approach to estimate small areas in survey sampling. In section 3 we study the behavior

of the prior specification and posterior models for our proposal. In section 4 the potential of our proposal is illustrated in a popular example of the batting averages for 18 players called the "Clemente Problem" given in Efron & Morris (1975). Some closing concluding remarks are presented in Section 5.

2 Objective Robust Bayesian approach to binary data

A simple motivating example

We introduce the robust Bayesian point of view in survey sampling using an example of simple random sampling for a finite population. We use the notation given in Little & Zheng (2007) and Gelman, Carlin, Stern & Rubin (2004). Consider a finite population $U = \{1, ..., N\}$ and $y = (y_1, ...y_N)$ denote the values of a variable in the population. Consider the marginal distribution of y over the prior distribution with parameter θ:

$$p(y) = \int \prod_{i=1}^{N} p(y_i|\theta)p(\theta)d\theta \tag{1}$$

to draw a simple random sample of size n. For this problem, the estimated of interest is the average of the finite-population $\bar{y}$:

$$\bar{y} = \frac{n}{N}\bar{y}_{obs} + \frac{N-n}{N}\bar{y}_{mis} \tag{2}$$

where $\bar{y}_{obs}$ and $\bar{y}_{mis}$ are the averages of the observed and missing y_i's respectively. Assume that $y_i|\theta$ has a normal distribution where $\mu = E(yi|\theta)$ and $\sigma^2 = V(yi|\theta)$ are the expectation and variance of the y_i's. If $N - n$ is large $p(\bar{y}_{mis}|\theta) \approx N(\bar{y}_{mis}|\mu, \sigma^2/(N-n))$ which denotes a normal density on $\bar{y}_{mis}$ with mean μ and variance $\sigma^2/(N-n)$, respectively. Using the standard noninformative prior distribution for $p(\mu, log(\sigma^2))$, we have the exact result $\bar{y}|y_{obs} \sim t_{n-1}(\bar{y}_{obs}, s^2_{obs}(1/n - 1/N))$ where t_{n-1} denotes the student-t distribution with $n-1$ with degree of freedom (see Gelman et al. (2004)). Suppose that prior information is available for the location parameter of μ, then we can use $\mu \sim N(\theta, \sigma^2)$, obtaining:

$$\bar{y}|y_{obs} \sim t_{n-1}((\theta + n\bar{y}_{obs})/(n+1), (N-n)(N+1)s^2_{obs}/(N^2(n+1))). \tag{3}$$

We can see in (3) that the mean in is a convex combination of the prior expectation, θ, and the data average, $\bar{y}_{obs}$, and thus the prior has *unbounded influence*. For example, as the location prior/data conflict $|\theta - \bar{y}_{obs}|$ grows, so does $|\theta_n - \bar{y}_{obs}|$ and without bound. These considerations motivate the interest in non-conjugate models for Bayesian analysis of survey sampling, and more generally motivate the use of objective robust Bayesian heavy-tailed priors. We find in the literature different proposals about robust priors. For example, in Dawid (1973), O'Hagan (1979), Evans & Moshonov (2006), Gelman, Jakulin, Pittau & Su (2008) and Pericchi, Sanso & Smith (1993), where robust priors for location parameters are studied. However, our proposal is about the natural parameter of the Binomial likelihood in the exponential family (not location

parameter) with objective robust priors recently known.

Additionally, most surveys have binary data and the Binomial likelihood is appropriate to model the data. In these surveys is very common that estimates of proportions are desired for subpopulations (domains). However, sometimes the sample size for a given domain is very small and it is necessary to provide a useful estimate for this small areas. Let θ_i be the true proportion of having a particular characteristic in the small area and the data $\{y_{ij}, i = 1, 2, ..., m, j = 1, ..., n_i\}$, where y_{ij} is the value of the jth unit belonging to the ith area. The data can be reduced to $y_i = \sum_{j=1}^{ni} y_{ij} \sim \text{Binomial}(n_i, \theta_i)$; $i = 1, ..., m$. One approach for this problem is to use the usual conjugate Beta prior with parameters a and b to make estimation for each small area with the conjugate analysis, but the influence of the prior information in the Beta prior could be very high when prior and likelihood information are in conflict (see Fúquene et al. (2009)). On the other hand, the Binomial likelihood in the exponential family form is:

$$p(y_i|\lambda_i) \propto \exp\{y_i\lambda_i - n_i \log(1 + e^{\lambda_i})\}, \tag{4}$$

where the natural parameter is the log-odds $\lambda_i = \log(\theta_i/(1 - \theta_i))$, $-\infty < \lambda_i < \infty$. The posterior expectation and variance using (4) and a Beta(a, b) prior, after of the transformation the parameter θ to log-odds, are $E_{BB}(\lambda) = \Psi(a) - \Psi(b)$ and $V_{BB}(\lambda) = \Psi^{'}(a) + \Psi^{'}(b)$ where $\Psi(\cdot)$ is Digamma function and $\Psi^{'}(\cdot)$ is Trigamma function. Changing the hyperparameters a, b, the expectation can be changed without a bound, so the influence of the prior mean is unbounded. In order to obtain robust analysis of binary data for small area estimation we can use the three following models:

2.1 First Robust Model

Fúquene et al. (2009) present a novel result, The Polynomial Tails Comparison Theorem, which gives easy-to-check conditions to ensure prior robustness for the natural parameter in the exponential family. The authors considered a Cauchy prior for the Binomial likelihood, where the conditions of their theorem are available. For this reason the first robust analysis for binomial data is a Cauchy prior for the natural parameter λ_i:

$$p_C(\lambda_i) = \frac{\tau_i}{\pi[\tau_i^2 + (\lambda_i - \mu_i)^2]}; \qquad \tau_i > 0 \quad -\infty < \mu_i < \infty \tag{5}$$

in order to achieve robustness with respect to the prior. Hence, the first robust model is

$$\begin{aligned} y_i &\sim \text{Binomial}(n_i, \lambda_i), \\ \lambda_i &\sim \text{Cauchy}(\mu_i, \tau_i), \end{aligned} \tag{6}$$

where μ_i and τ_i are the parameters of localization and scale of the Cauchy prior. If prior information is available for all small areas we can use it in the cauchy prior having the prior information in the log-odds scale as follows: $\mu_i = \Psi(x_i + a_i) - \Psi(n_i - x_i + b_i)$ and $\tau_i = (\Psi^{'}(x_i + a_i) + \Psi^{'}(n_i - x_i + b_i))^{1/2}$ where x_i is the auxiliary variable (i.e. prior information) for the ith small area.

2.2 Second Robust Hierarchical Model

On the other hand, Fúquene et al. (2011) propose to use the Beta Distribution scaled of the Second Kind, (or Beta 2 scaled distribution) for the square scale parameters in dynamic linear models for modelling outliers and structural breaks. In order to estimate proportions in small areas, we use this prior for the square scale parameter in hierarchical models. The scaled Beta2 prior for the square scale is the following:

$$p(\tau_i^2) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \frac{1}{\beta} \frac{\left(\frac{\tau_i}{\beta}\right)^{p-1}}{\left(1+\frac{\tau_i}{\beta}\right)^{p+q}}. \quad \tau > 0. \tag{7}$$

For precisions $\phi_i = 1/\tau_i$, we assign the scaled Beta 2 as

$$p(\phi_i) = \frac{\Gamma(q+p)}{\Gamma(q)\Gamma(p)} \beta \frac{(\beta\phi_i)^{q-1}}{(1+\beta\phi_i)^{p+q}}; \quad \phi_i > 0, \tag{8}$$

The marginal of the location parameters of, a Beta 2 scale density for $p = q = 1$ for the square scale parameters coupled with a Cauchy prior for the location, is a novel heavy tailed prior (see Fúquene et al. (2011)):

$$p(\lambda_i) = \frac{1}{2\sqrt{\beta_i}\left(1+\frac{|\lambda_i - M|}{\sqrt{\beta_i}}\right)^2} \tag{9}$$

(see proof in the appendix). The Fúquene et al. (2011) prior has the following qualities: 1) is proper. 2) has tails heavier than the Cauchy prior. Therefore, we consider the robust hierarchical model as follow:

$$y_i \sim \text{Binomial}(n_i, \lambda_i), \tag{10}$$
$$\lambda_i|\phi \sim \text{Cauchy}(M, \phi_i), \tag{11}$$
$$\phi_i \sim \text{Beta2}(1, 1, 1/\beta),$$

where $\text{Beta2}(1, 1, 1/\beta)$ is an independent Scaled Beta2 for the square scale in each small area and M is taken as the general mean obtained using the prior information for all small areas.

In order to compare the Cauchy, Normal and Fuquene et al (2011) prior we make a match of the quartiles equal ± 1. Therefore, the scale for the Normal is 1.47 and for both Cauchy and Fúquene et al. (2011) prior the scale is 1. Figures 1 and 2 display that the Fuquene et al (2011) prior has tails heavier than the Cauchy prior.

Figure 1: Comparison of the Student-t-Beta2(1,1,1,1) ,Cauchy(0,1), Normal(0,2.19) priors.

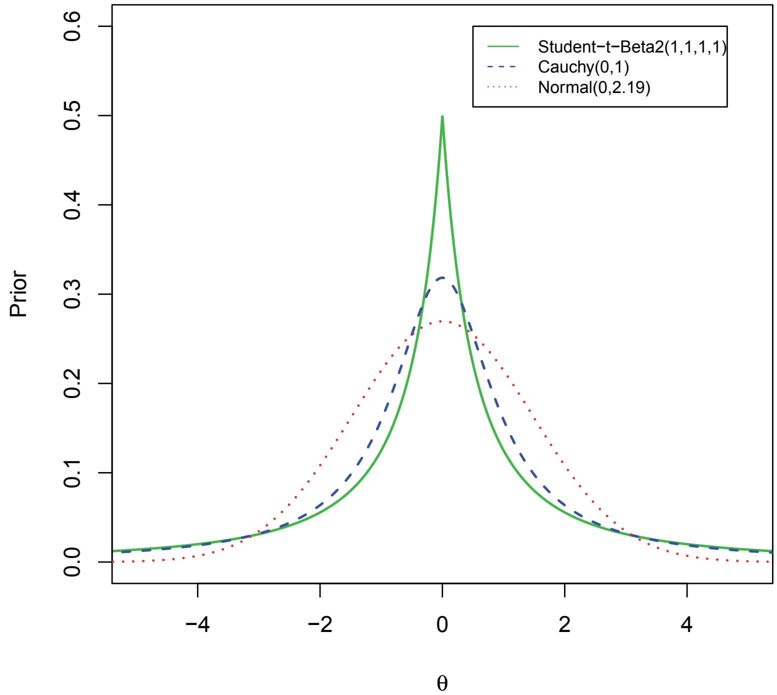


Figure 2: Comparison of the tails of the Student-t-Beta2(1,1,1,1) ,Cauchy(0,1), Normal(0,2.19) priors.

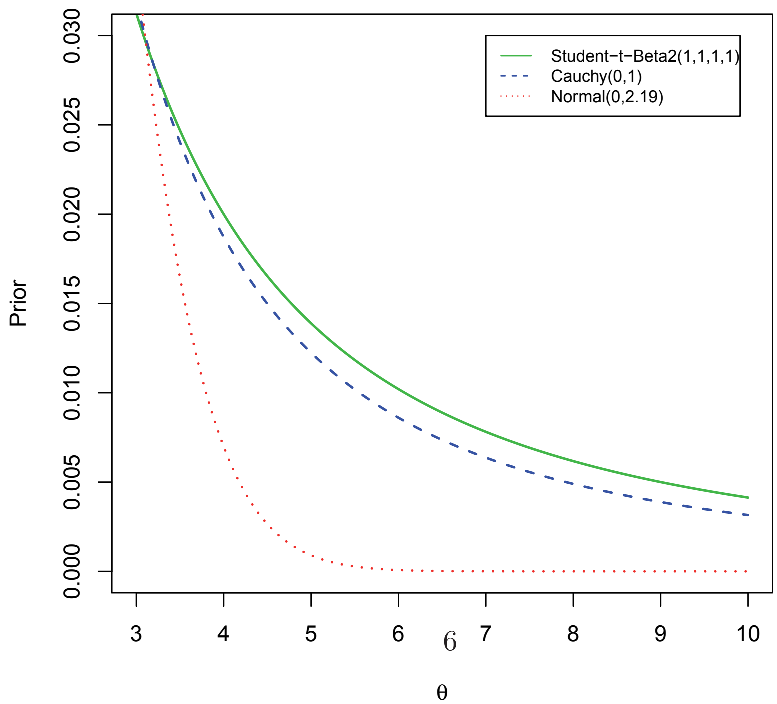


Figure 1 displays the Cauchy$(0, 1)$, Normal$(0, 2.19)$, Horseshoe and Cauchy-Beta$(0, 1, 1)$ priors. The Horseshoe has heavy tails as the Cauchy prior and the Cauchy-Beta$(0, 1, 1)$ has tails even heavier than these.

3 Illustration with the three approaches

We compare the three approaches observing the posterior predictive mean and variance as functions of the discrepancy between the MLE and prior location. For the Cauchy prior we use the R (R Development Core Team (2011)) package named ClinicalRobustPriors (see Fuquene (2009)) to compute probabilities and figures for the prior, likelihood and posterior models. On the other hand, for the Fúquene et al. (2011) prior the BRugs package is used(see Thomas, O'Hara, Ligges & Sturtz (2006)). The MLE for the natural parameter of the Binomial likelihood is kept fixed at $\log(\overline{y}_i/(1-\overline{y}_i)) = 0$ and the prior location is moved to create a conflict between data and prior.

From Figure 2, we can see that with Cauchy and Fuquene et al (2011) priors the estimation of the posterior predictive mean tends to the MLE. This behavior is expected for Bayesian robustness; therefore, these priors are robust for the Binomial likelihood in the exponential family. In other words the influence of the prior is bounded. Figure 3 displays the posterior variance, we can see that with the Cauchy and Fúquene et al. (2011) priors the posterior variance is not monotonic in the conflict between the MLE and prior location.

Figure 3: Posterior Predictive Mean using the Cauchy and Student-t-Beta2(1,1,1,1) priors.

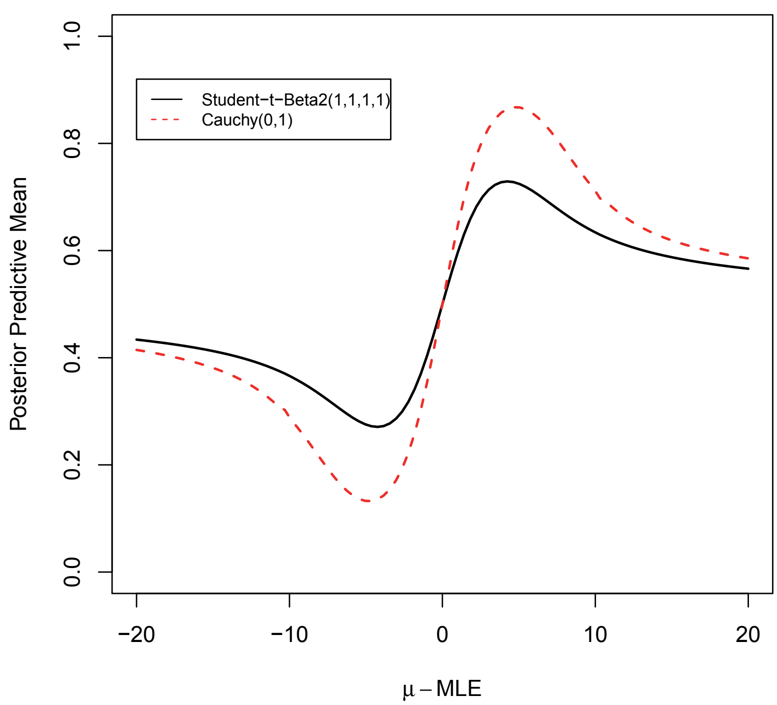
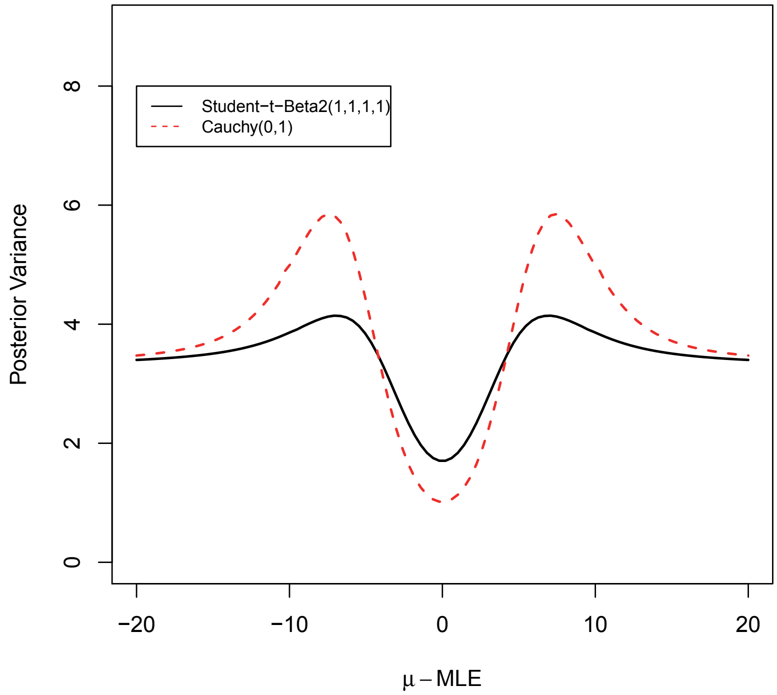


Figure 4: Posterior variance using the Cauchy and Student-t-Beta2(1,1,1,1) priors.



4 Example: "The Clemente problem"

In this section we apply our proposal in an historical example given in Efron & Morris (1975). This data set has been explored by different authors including Morris (1983), Gelman, Carlin, Stern & Rubin (1995), Datta & P. (2000), Rao (2003), Jiang & Lahiri (2006) and Pericchi & Perez (2010). Efron & Morris (1975) consider the problem of estimating the batting averages of 18 baseball players in the 1970 season; since the batting averages are small, each player is considered as a small area. Although this is an estimation proportion problem, they used hierarchical Normal/Normal model transforming the data to obtain the estimates of the proportions. We consider this estimation problem in a different way using two auxiliary variables given in Gelman et al. (1995) such as the batting average of each player in the previous season of 1969 and the number of times at bat in that season. For the Binomial likelihood the sample size is $n_i = 45$ the number of times at bat and y_i the number of hits among n_i for the ith player. θ_i is the true know batting average for the 1970 season.

Table 1: Estimates of the posterior predictive mean and relative root mean squared error for the conjugate and non-conjugate analysis from the battering averages for 18 players in the 1970 season using two variables of 1969 season as auxiliary information

Player θ_i	$\hat{\theta}_i$ (MLE) $\frac{\exp(\mu_i)}{1+\exp(\mu_i)}$	B/B RRMSE	C/B RRMSE	N/B-Inv RRMSE	Student-t-Beta2(1,1,1,$\boldsymbol{\beta_i}$) RRMSE
Clemente	**0.400**	**0.314**	**0.317**	**0.289**	**0.372**
0.352	**0.314**	**0.108**	**0.099**	**0.179**	**0.057**
F. Robinson	0.378	0.303	0.309	0.284	0.353
0.306	0.303	0.010	0.010	0.072	0.154
Munson	0.178	0.229	0.219	0.234	0.196
0.302	0.256	0.242	0.275	0.225	0.351
Scott	0.222	0.249	0.250	0.248	0.231
0.296	0.250	0.159	0.155	0.162	0.220
F. Howard	0.356	0.276	0.276	0.279	0.334
0.283	0.275	0.025	0.025	0.014	0.180
Campaner	0.200	0.263	0.265	0.244	0.213
0.279	0.264	0.057	0.050	0.125	0.237
Spencer	0.311	0.247	0.256	0.269	0.298
0.276	0.246	0.105	0.072	0.025	0.080
Berry	0.311	0.250	0.261	0.267	0.297
0.274	0.244	0.088	0.047	0.026	0.084
Swoboda	0.244	0.281	0.271	0.252	0.246
0.267	0.281	0.052	0.015	0.056	0.079
Kessinger	0.289	0.249	0.247	0.263	0.278
0.266	0.248	0.064	0.071	0.011	0.045
E Rodriguez	0.222	0.254	0.251	0.245	0.233
0.261	0.255	0.027	0.038	0.061	0.107
Willians	0.222	0.256	0.243	0.248	0.230
0.258	0.257	0.008	0.058	0.039	0.109
Unser	0.222	0.269	0.266	0.248	0.233
0.251	0.271	0.072	0.060	0.012	0.072
Johnstone	0.333	0.258	0.268	0.272	0.315
0.238	0.255	0.084	0.126	0.143	0.324
Santo	0.244	0.244	0.240	0.256	0.246
0.233	0.244	0.047	0.030	0.099	0.056
Petrocelli	0.222	0.232	0.221	0.248	0.231
0.225	0.234	0.031	0.018	0.102	0.027
Alvarado	**0.267**	**0.188**	**0.224**	**0.258**	**0.263**
0.224	**0.118**	**0.161**	**0.000**	**0.152**	**0.174**
Alvis	0.156	0.248	0.233	0.233	0.179
0.183	0.249	0.355	0.273	0.273	0.022

From Table 1 we observe two different types of outliers. The first one is an outlier with respect to the prior information in the small area. The second type is a small area resulting in conflict with the other small areas. The first outlier is Alvarado because the average during the first 45 at-bats at 1970 (0.224) is much better than his previous batting average at 1969 (0.118). For the first outlier we use a Cauchy prior for the small area. The second outlier is the player Roberto Clement who undoubtedly was an extremely good hitter not only during the 1969 season but also during many seasons. For the second we can use Fuquene et al (2011) prior. In this example we can see that for the Cauchy/Binomial (C/B) model the estimation of the average for Alvarado has a relative root mean squared error, RRMSE $= \sqrt{\text{MSE}}/(\text{true value})$, equal to zero. For this outlier the conflict between prior and likelihood is equal to $|0.267 - 0.188| = 0.159$ and the estimation is equal to the parameter (0.224). However, using the Beta/Binomial (B/B) conjugate model for the small area the RRMSE is equal to 0.161 and with this conjugate prior the influence of the auxiliary information is very high. On the other hand, using the Fúquene et al. (2011) prior the estimation for the player Roberto Clemente (≈ 0.372 and 0.373) is very close to the true value (0.352), and the influence of the mean of the prior information, $e^{-1.079}/(1+e^{-1.079}) = 0.253$, is discounted. However, using the Normal/Binomial and a non-informative Inverted Gamma for the square of the scale (N/B-Inv-Gamma), the estimation (≈ 0.289) is very influenced for the prior information given for all small areas. Finally, we can see that when there is no conflict between the auxiliary information within the small areas, the Cauchy prior provides approximately the same results than the conjugate analysis. This fact is a very important quality for using a Cauchy prior as a objective default prior.

5 Concluding remarks

In recent years different methodologies using objective robust priors have been proposed. Examples are methodologies for important areas such Clinical trials (Fúquene et al. (2009)), Quality Control (Bayarri & Garcia-Donato (2005)) and Genetics (Consonni & Moreno (2011)) to mention only some of them. However in survey sampling there is no a clear proposal. An objective robust Bayesian methodology could be more acceptable for both practitioners involve in survey sampling and agencies making these surveys in order to eliminate antipathy towards methods that involve subjective elements or assumptions . We propose to use the Cauchy and Fuquene et al (2011) priors as default (and robust) priors in the estimation of proportions on small areas: 1) For the estimation of proportions when there is a conflict or not between the auxiliary information (prior information) and the data within the small area the best alternative, as a objective default prior for Bayesian robustness, is a Cauchy prior. 2) The second type of outlier is given when one small area is in conflict with the rest of small areas. In this case we need robustness with respect to the prior information for all small areas. Therefore, we recommend to use the Fúquene et al. (2011) prior as an objective default prior in order to obtain robustness with respect to the prior information. 3) MCMC simulations for our proposal can be made easily using either of the ClinicalRobustPriors or BRugs R packages 4) In a future work these approaches can be explored in important designs such as stratified and multistage.

A Appendix

Fuquene et al (2011) found a new class of hypergeometric heavy tailed priors. They consider the Student-t density coupled with the scaled Beta2 prior to the square of the scale as follow:

Result: Let $\theta \sim$ Student-t(μ, τ, υ) where υ are the degrees of freedom, μ the location and τ the scale of the Student-t density:

$$\pi(\theta|\tau^2) = \frac{k_1}{\tau}\left(1 + \frac{1}{\upsilon}\left(\frac{\theta-\mu}{\tau}\right)^2\right)^{-(\upsilon+1)/2}, \quad \upsilon > 0, -\infty < \mu < \infty, -\infty < \theta < \infty, \tag{12}$$

where $k_1 = \dfrac{\Gamma((\upsilon+1)/2)}{\Gamma(\upsilon/2)\sqrt{\upsilon\pi}}$. Therefore

$$\pi(\theta) = \begin{cases} k\beta^q\nu/(\theta-\mu)^{q+1/2}2F1(p+q, q+1/2, (\upsilon+1)/2+p+q, 1-\beta\nu/(\theta-\mu)^2) & \text{if} \quad \theta \neq \mu, \\ \\ k_1\mathrm{Be}(p-1/2, q+1/2)/(\beta^{1/2}\mathrm{Be}(p,q)) & \text{if} \quad \theta = \mu. \end{cases}$$

with $k = k_1\mathrm{Be}(q+1/2, p+\upsilon/2)/\mathrm{Be}(p,q)$. Where $Be(a,b)$ denotes the beta function and $2F1(a,b,c,z)$ denotes the hypergeometric function.

We can find (9) using the identities 15.3.3 and 15.1.13 for $\theta \neq \mu$ and $p = q = \upsilon = 1$ of Abramowitz & Stegun (1970) as follow:

$$\begin{aligned} 2F1(2, 3/2, 3, 1-\beta\nu/(\theta-\mu)^2) &= ((\theta-\mu)/\beta^{1/2})2F1(1, 3/2, 3, 1-\beta/(\theta-\mu)^2) &\quad (13)\\ &= 4((\theta-\mu)/\beta^{1/2})(1+|\beta^{1/2}/(\theta-\mu)|)^{-2} &\quad (14) \end{aligned}$$

and $k = k_1\mathrm{Be}(3/2, 3/2)/\mathrm{Be}(1,1) = 1/8$, therefore:

$$\pi(\theta) = \beta^{1/2}(\theta-\mu)^{-1/2}(1+|\beta^{1/2}/(\theta-\mu)|)^{-2}/2 \tag{15}$$

for $\theta = \mu$

$$\pi(\theta) = k_1\mathrm{Be}(1/2, 3/2)/(\beta^{1/2}\mathrm{Be}(1,1)) = \frac{1}{2\beta^{1/2}}. \tag{16}$$

References

Abramowitz, M. & Stegun, I. (1970), *Handbook of Mathematical Functions. National Bureau of Standards*, Vol. 46, Applied Mathematics Series.

Bayarri, M. J. & Garcia-Donato, G. (2005), 'A bayesian sequential look at u-control charts', *Journal of the American Statistical Association* **47-2**, 142–151.

Consonni, G. & Moreno, S. E. (2011), 'Testing hardyweinberg equilibrium: An objective bayesian analysis', *Statistics in Medicine* **30**, 67–74.

Datta, G. S. & P., L. (2000), 'A unified measure of uncertainty of estimated best linear unbiased predictors in small area estimation problems', *Statistica Sinica* **10**, 613–627.

Dawid, A. P. (1973), 'Posterior expectations for large observations', *Biometrika* **60**, 664–667.

Efron, B. & Morris, C. (1975), 'Data analysis using steins estimator and its generalizations', *Journal of the American Statistical Association* **70**, 311–319.

Evans, M. & Moshonov, H. (2006), 'Checking for prior-data conflict', *Bayesian Analysis* **1**, 893–914.

Fuquene, J. A. (2009), *ClinicalRobustPriors: Robust Bayesian Priors in Clinical Trials: An R Package for Practitioners*. R package version 2.1-2.
*http://CRAN.R-project.org/package=ClinicalRobustPriors

Fúquene, J. A., Cook, J. D. & Pericchi, L. R. (2009), 'A case for robust bayesian priors with applications to clinical trials', *Bayesian Analysis* **4**, 817–846.

Fúquene, J. A., Perez, M. E. & Pericchi, L. R. (2011), 'Modelling outliers and structural breaks in dynamic linear models with a novel use of a heavy tailed prior for the variances: An alternative to the inverted gamma', *Technical Report arXiv:1107.1811v1* .
*http://arxiv.org/abs/1107.1811v1

Gelman, A., Carlin, J., Stern, H. & Rubin, D. (1995), *Bayesian Data Analysis*, Chapman and Hall, London.

Gelman, A., Carlin, J., Stern, H. & Rubin, D. (2004), *Bayesian Data Analysis*, 2 edn, Chapman and Hall, London.

Gelman, A., Jakulin, A., Pittau, M. G. & Su, Y.-S. (2008), 'A weakly informative default prior distribution for logistic and other regression models', *Annals of Applied Statistics* **2**, 1360–1383.

Jiang, J. & Lahiri, P. (2001), 'Empirical best prediction for small area inference with binary data', *Annals of the Institute of Statistical Mathematics* **53**, 217–243.

Jiang, J. & Lahiri, P. (2006), 'Mixed model prediction and small area estimation', *Test* **15**, 1–96.

Jonhnson, N., Kotz, S. & Balakrishnan, N. (1995), *Continuous Univariate Distributions*, Vol. 8, 2 edn, Wiley, New York.

Little, R. J. A. & Zheng, H. (1980), The analysis of census undercount from a postenumeration survey, *in* 'Proceedings of the Conference on Census Undercount', pp. 88–94.

Little, R. J. A. & Zheng, H. (2007), The bayesian approach to the analysis of finite population survey, *in* 'Bayesian Statistics', Vol. 8, D. Heckerman, A. F. M Smith and M. West, Oxford University Press, pp. 283–302.

Morris, C. N. (1983), 'Parametric empirical bayes inference: theory and applications', *Journal of the American Statistical Association* **78**, 47–65.

O'Hagan, A. (1979), 'On outlier rejection phenomena in Bayes inference', *Journal of the Royal Statistics Society* **B, 41**, 358–367.

Pericchi, L. R. & Perez, M. E. (2010), 'Limiting the shrinkage for the exceptional by objective robust bayesian analysis', *Technical Report Centro de Bioestadstica y Bioinformtica* .
*http://repositorio.upr.edu

Pericchi, L. R., Sanso, B. & Smith, A. F. M. (1993), 'Posterior cumulant relationships in bayesian inference involving the exponential family', *Journal of the American Statistical Association* **88**, 1419–1426.

R Development Core Team (2011), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
*http://www.R-project.org/

Rao, J. N. K. (2003), *Small Area Estimation*, Wiley Seies in Survey Methodology, Hoboken, NJ Wiley-Interscience, John Wiley and Sons.

Stroud, T. W. F. (1991), 'Hierarchical baxes predictive means and variances with application to sample survey inference', *Communications in Statistics-Theory and Methods* **20**, 13–36.

Stroud, T. W. F. (1994), 'Bayesian analysis of binary survey data', *Canadian Journal of Statistics* **22**, 33–45.

Thomas, A., O'Hara, B., Ligges, U. & Sturtz, S. (2006), 'Making bugs open', *R News* **6**(1), 12–17.
*http://cran.r-project.org/doc/Rnews/